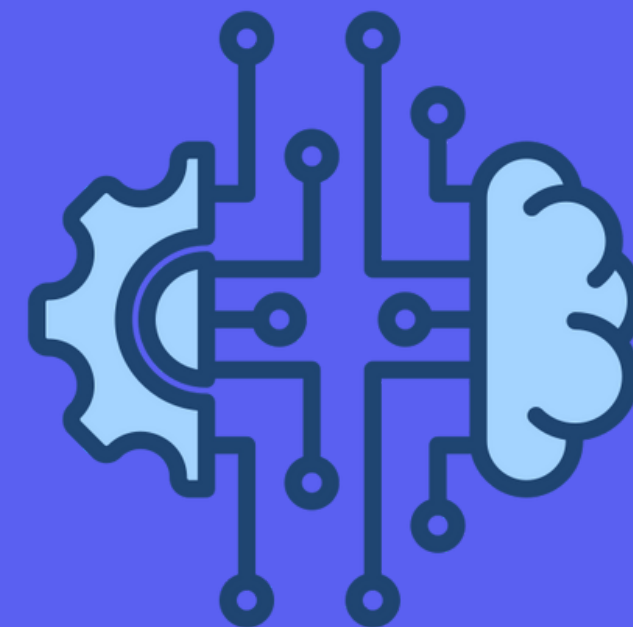




31/03/2023

Orientador: Prof. Dr. Daniel
Sabino Amorim de Araujo

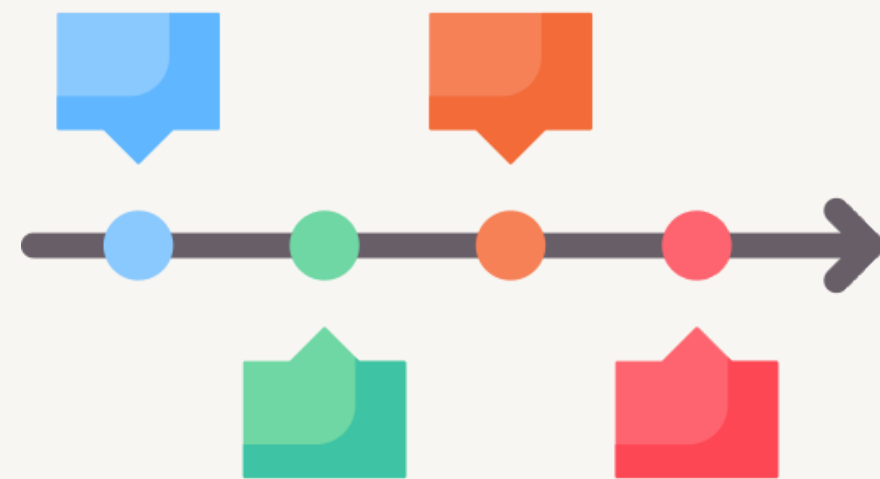
Coorientador: David Montalv o
Junior



Desenvolvimento de modelo de Machine Learning para prioriza  o de peti  es iniciais em processos judiciais

Jo  o Paulo de Oliveira C  mara Fernandes

Cronograma de Apresentação



- Introdução

- Problematização

- Objetivo

- Metodologia

- Resultados

- Considerações Finais

Introdução



O Judiciário julgou 26,9 milhões de processos em 2021, o que significou uma alta de 11,1% em relação aos números de casos solucionados do ano anterior (Justiça em Números, 2022)



Foi registrado 7,7 milhões de novas ações em 2021, o que representa 10,4% de crescimento



Investimento em um alto valor em material humano para equilibrar o fluxo de entradas e saídas de processos dentro da justiça



A justiça brasileira já é frequentemente criticada por seus altos gastos

Introdução



Adoção de tecnologias e processos mais eficientes, como o uso de inteligência artificial (IA) e automação de processos.



O desenvolvimento de IA no âmbito jurídico está sendo cada vez mais difundido



Até o meio do ano de 2022 existiam 111 projetos sendo desenvolvidos no Poder Judiciário que envolviam IA



Aumento de 171% em relação ao mesmo período do ano anterior (CNJ)

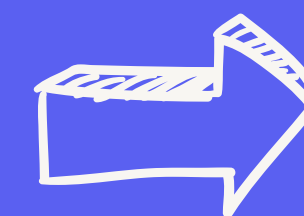
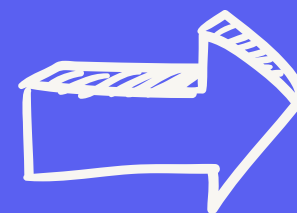
Problema



Sobrecarga de processos no Poder Judiciário brasileiro



Processos com necessidades de prioridade diferentes



NORMAL

PREFERENCIAL

URGENTE

Objetivos



Automatizar a tarefa de distribuir o processo para o setor correto



Proporcionar agilidade no andamento de processos que necessitam de urgência



Otimizar tempo do Servidor

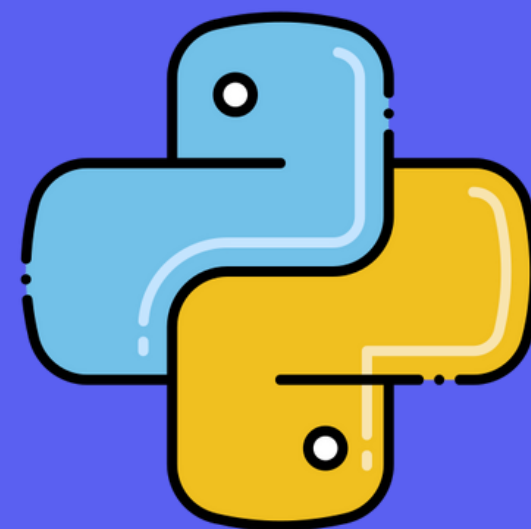


Parcela de ajuda para mitigar o congestionamento crescente na Justiça

Metodologia



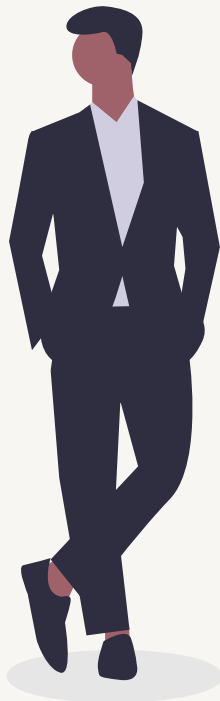
Ferramentas



Google Colaboratory

Metodologia

Aquisição dos dados



2022 2021 2020 2019 Jan Feb Mar 2018 2017 2016 2015 Jul Ago Set

Colegiado: Serventia:

CnONCrim2 - Casos Novos Originários no 2º Grau Não

Processos Preferência Legal

| Mês/Ano | Preferência Legal | Serv |
|----------|-------------------|----------|
| Jan/2022 | 0 | Roberto |
| Jan/2022 | 0 | Paulo R |
| Jan/2022 | 0 | Leonard |
| Jan/2022 | 0 | Leonard |
| Jan/2022 | 0 | Vladimir |
| Jan/2022 | 0 | Manoel |
| Jan/2022 | 0 | Paulo M |
| Jan/2022 | 0 | Rubens |
| Jan/2022 | 0 | Rogério |
| Jan/2022 | 0 | Roberto |
| Jan/2022 | 0 | Manoel |
| Jan/2022 | 0 | Rubens |
| Jan/2022 | 0 | Rubens |
| Jan/2022 | 0 | Carlos F |
| Jan/2022 | 0 | Roberto |
| Jan/2022 | 0 | Roberto |
| Jan/2022 | 0 | Cid Mar |
| Jan/2022 | 0 | Leonard |
| Jan/2022 | 0 | Leonard |

Processos

| Preferência Legal | Mês/Ano | Nú |
|-------------------|----------|--------|
| | Jan/2022 | 080001 |
| Idoso(a) | Jan/2022 | 080002 |
| Idoso(a);Pes... | Jan/2022 | 080002 |
| Idoso(a) | Jan/2022 | 080004 |
| | Jan/2022 | 080004 |
| | Jan/2022 | 080005 |
| | Jan/2022 | 080005 |
| | Jan/2022 | 080008 |
| Idoso(a) | Jan/2022 | 080008 |
| | Jan/2022 | 080008 |
| | Jan/2022 | 080009 |
| Idoso(a) | Jan/2022 | 080009 |
| Idoso(a) | Jan/2022 | 080009 |
| | Jan/2022 | 080009 |
| | Jan/2022 | 080009 |
| Idoso(a) | Jan/2022 | 080011 |

Metodologia



Aquisição dos dados



Metodologia



Aquisição dos dados



Metodologia



Análise exploratória dos Dados

Dados do TRF5

| | %ID_PROCESSO_TRF | Sistema | Data Protocolo | Data Início Processo | Número Processo | Eletrônico? | Eletrônico/Físico | Instância Processo | Data Prim Distribuição | Data Últ Distribuição | SJ Processo | Flag Preferência Legal | Preferência Legal |
|---|------------------|---------|---------------------|----------------------|---------------------------|-------------|-------------------|--------------------|------------------------|-----------------------|-------------|------------------------|-------------------|
| 0 | PJe AL 192169 | PJe | 14/02/2019 14:16:02 | 2/14/2019 1:39:12 PM | 0801602-50.2019.4.05.0000 | Sim | Eletrônico | 1º Grau | NaN | NaN | AL | NaN | NaN |
| 1 | PJe AL 193149 | PJe | 17/05/2022 15:29:29 | 5/17/2022 3:06:31 PM | 0805367-24.2022.4.05.0000 | Sim | Eletrônico | 1º Grau | NaN | NaN | AL | NaN | NaN |
| 2 | PJe AL 40036 | PJe | 24/01/2017 17:08:31 | 1/24/2017 5:08:30 PM | 0800736-69.2017.4.05.8000 | Sim | Eletrônico | 1º Grau | 42759.714259 | 42759.714259 | AL | NaN | NaN |



PJe, TEBAS, SEEU, PJeCNJ, Creta



AL, CE, PB, PE, RN, SE



Informações sobre processos e sua preferência legal



10.971.997 registros

Metodologia



Análise exploratória dos Dados

Dados do TRF5

➤ Flag Preferência Legal

```
df_trf5['Flag Preferência Legal'].unique()  
array([nan, 'Preferência Legal'], dtype=object)
```

Metodologia



Análise exploratória dos Dados

Dados do TRF5

➤ Preferência Legal

```
df['Preferência Legal'].value_counts()

Idoso(a) 652943
Prioridade de Tramitação 317187
Idoso(a) maior de 80 anos 66331
Deficiente Físico 38082
Pessoa com Deficiência 15664
Doença Terminal 13272
Grande Devedor 9774
Réu Preso 7901
Doença Terminal;Idoso(a) 3733
Deficiente Físico;Idoso(a) 1659
Idoso(a);Pessoa com Deficiência 1055
Idoso(a);Idoso(a) maior de 80 anos 505
Doença Terminal;Idoso(a) maior de 80 anos 416
Doença Terminal;Pessoa com Deficiência 200
Deficiente Físico;Doença Terminal 189
Pessoa em Situação de Rua 181
Idoso(a) maior de 80 anos;Pessoa com Deficiência 125
Preso/Acolhido/Internado 105
Idoso(a);Réu Preso 66
Doença Terminal;Idoso(a);Pessoa com Deficiência 47
Grande Devedor;Idoso(a) 43
Deficiente Físico;Doença Terminal;Idoso(a) 36
Doença Terminal;Idoso(a);Idoso(a) maior de 80 anos 23
```

```
set(lista_unica)

{'Deficiente Físico',
 'Doença Terminal',
 'Grande Devedor',
 'Idoso(a)',
 'Idoso(a) maior de 80 anos',
 'Pessoa com Deficiência',
 'Pessoa em Situação de Rua',
 'Pessoas com deficiência (art. 9º, VII, da Lei nº 13.146/2015)',
 'Preso/Acolhido/Internado',
 'Prioridade de Tramitação',
 'Réu Preso'}
```

Metodologia



Análise exploratória dos Dados

Pasta de Petições Iniciais

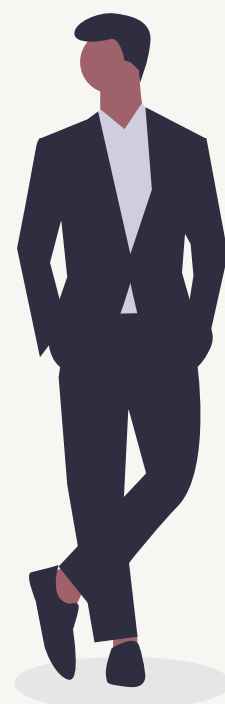


62200 documentos



Alguns documentos foram erroneamente incluídos pelos advogados

Metodologia



Análise exploratória dos Dados

Pasta de Petições Iniciais

Abzir ▾ ⬇ Baixar ... 0000639-13.2010.4.05.....pdf ⓘ Informações

Segue em anexo petição inicial e documentos digitalizados

Abzir ▾ ⬇ Baixar ... 0000318-61.2013.4.05.....pdf ⓘ Informações

Abzir ▾ ⬇ Baixar ... 0000077-21.2012.4.05.....pdf ⓘ Infor

☺☺☺

Digitalização para PJE do Processo 0000072-21.2012.4.05. [REDACTED] conforme Resolução Pleno T 12 de julho de 2017.

Metodologia



Análise exploratória dos Dados

Pasta de Petições Iniciais

0000003-44.2010.4.05.....pdf

Inf

TRIBUNAL REGIONAL FEDERAL DA 5ª REGIÃO
10ª VARA DA SEÇÃO JUDICIÁRIA DO RIO GRANDE DO NORTE
PROCESSO: 0000003-44.2010.4.05 EXEÇÃO FISCAL

| Polo ativo | Polo passivo |
|---------------------|--------------|
| A G EXEQUENTE | EXECUTADO |

Outros participantes
SEM REGISTROS

TERMO DE MIGRAÇÃO DE PROCESSO FÍSICO

Em cumprimento à Resolução Pleno nº 13, de 12 de julho de 2017, procedo à inclusão do presente feito no sistema de Processo Judicial Eletrônico - PJe, assegurando a conformidade da digitalização do processo físico às seguintes determinações da referida Resolução:

- utilização de funcionalidade específica para inserção das peças processuais digitalizadas;
- migração dos dados de autuação (número de registro do sistema físico, data de distribuição na Justiça Federal, etc.) e de movimentação;
- digitalização na íntegra do processo;
- compartimentação dos documentos conforme Anexo I da Resolução;
- cumprimento das regras de sigilo do processo e do sigilo do documento, quando for o caso;
- verificação de não se tratar de nenhum dos casos previstos no art. 5º da Resolução;
- intimação dos advogados no processo físico acerca da digitalização dos autos e da necessidade de cadastramento no PJe, quando não houver nenhum advogado habilitado no sistema;
- confirmação da unidade de Tecnologia da Informação competente de que há recursos de armazenamento suficientes para acomodar os feitos que serão incluídos no sistema PJe.

Mossoró, 15/02/2018

Metodologia

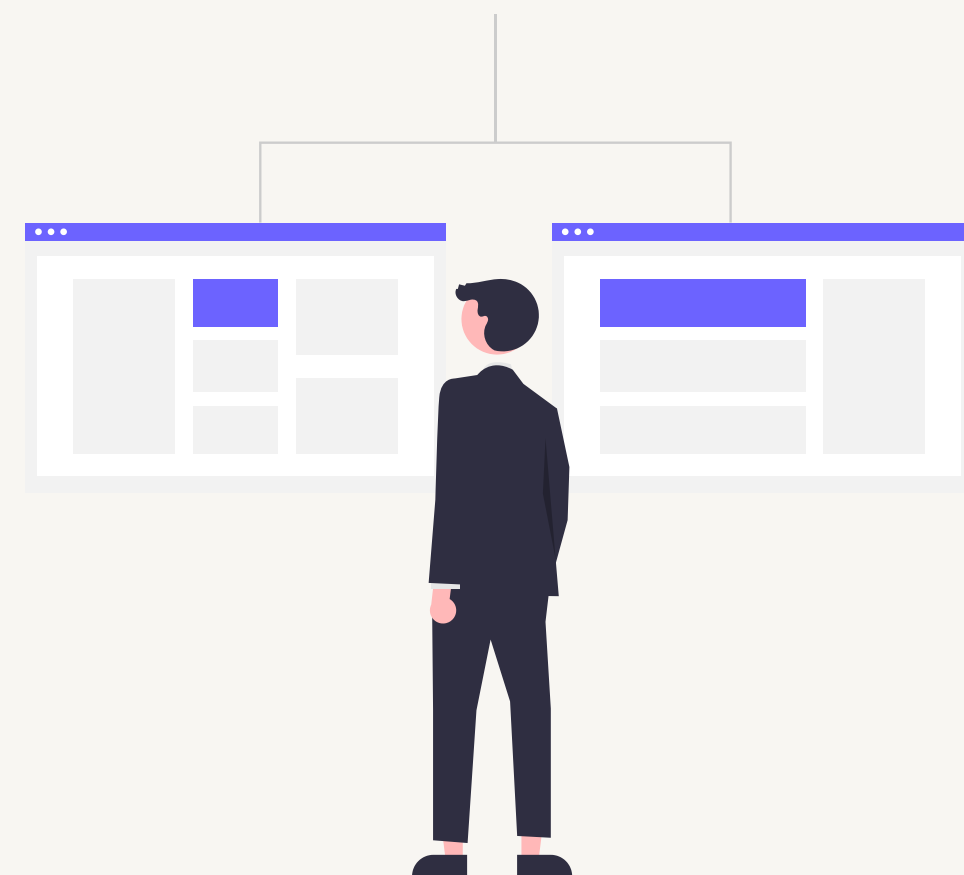


Limpeza dos Dados

Pasta de Petições Iniciais

- Removidos todos os documentos que eram termos de migração de processo físico
- Documentos entre 40 e 324 linhas
- Pelo menos 4000 caracteres
- 26136 documentos restantes

Metodologia



Cruzamento dos Dados



Número do processo é variável comum



Gerou dados apenas da Seção do RN

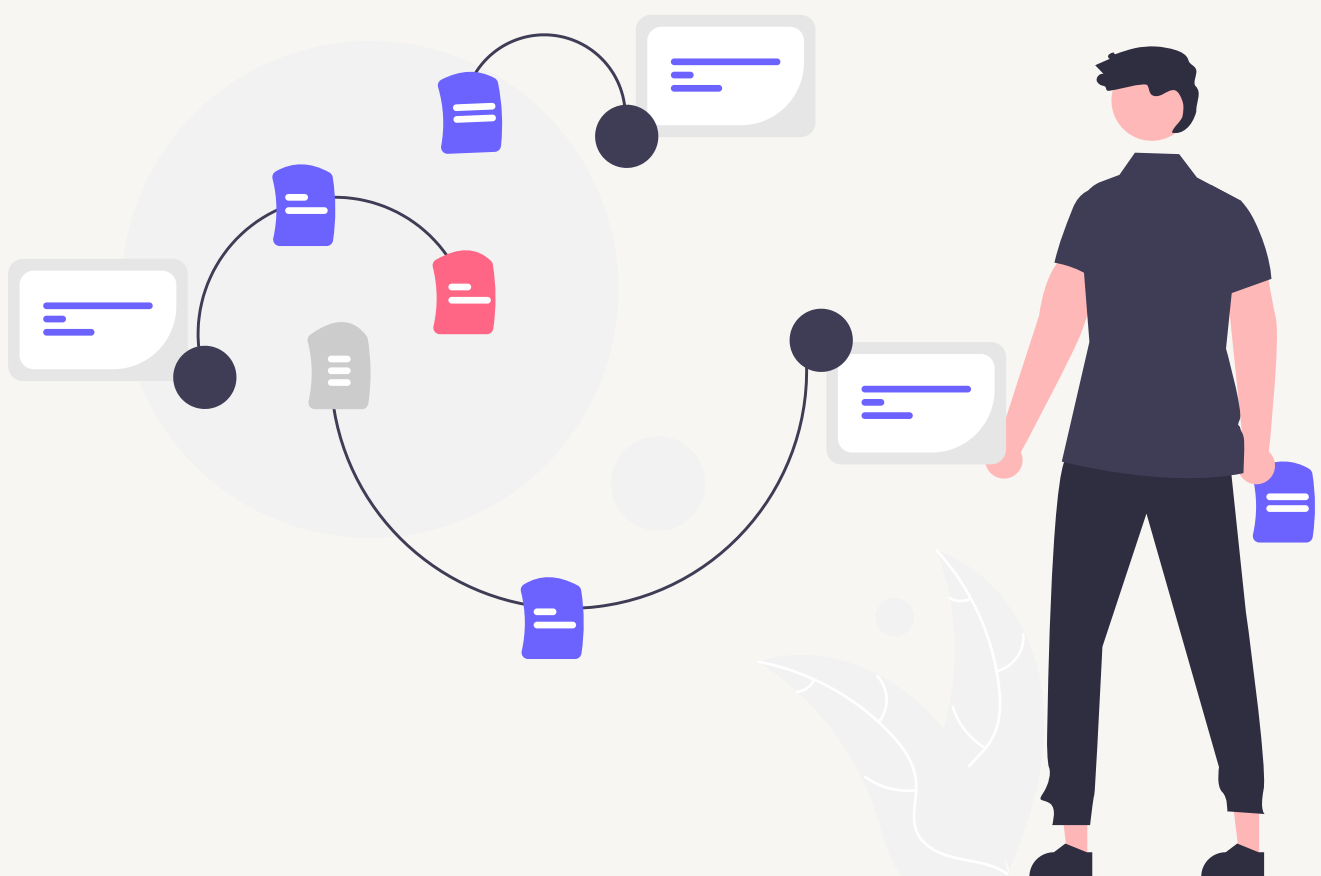
Metodologia



Pré-processamento do texto

- Stopwords
- Caracteres especiais
- Telefones, URLs, E-mails e símbolos de moeda
- Cabeçalho e rodapé
- Stematização

Metodologia



Treinamento dos modelos

- Identificar e rotular processos normais e preferenciais
- Identificar e rotular processos preferenciais e urgentes

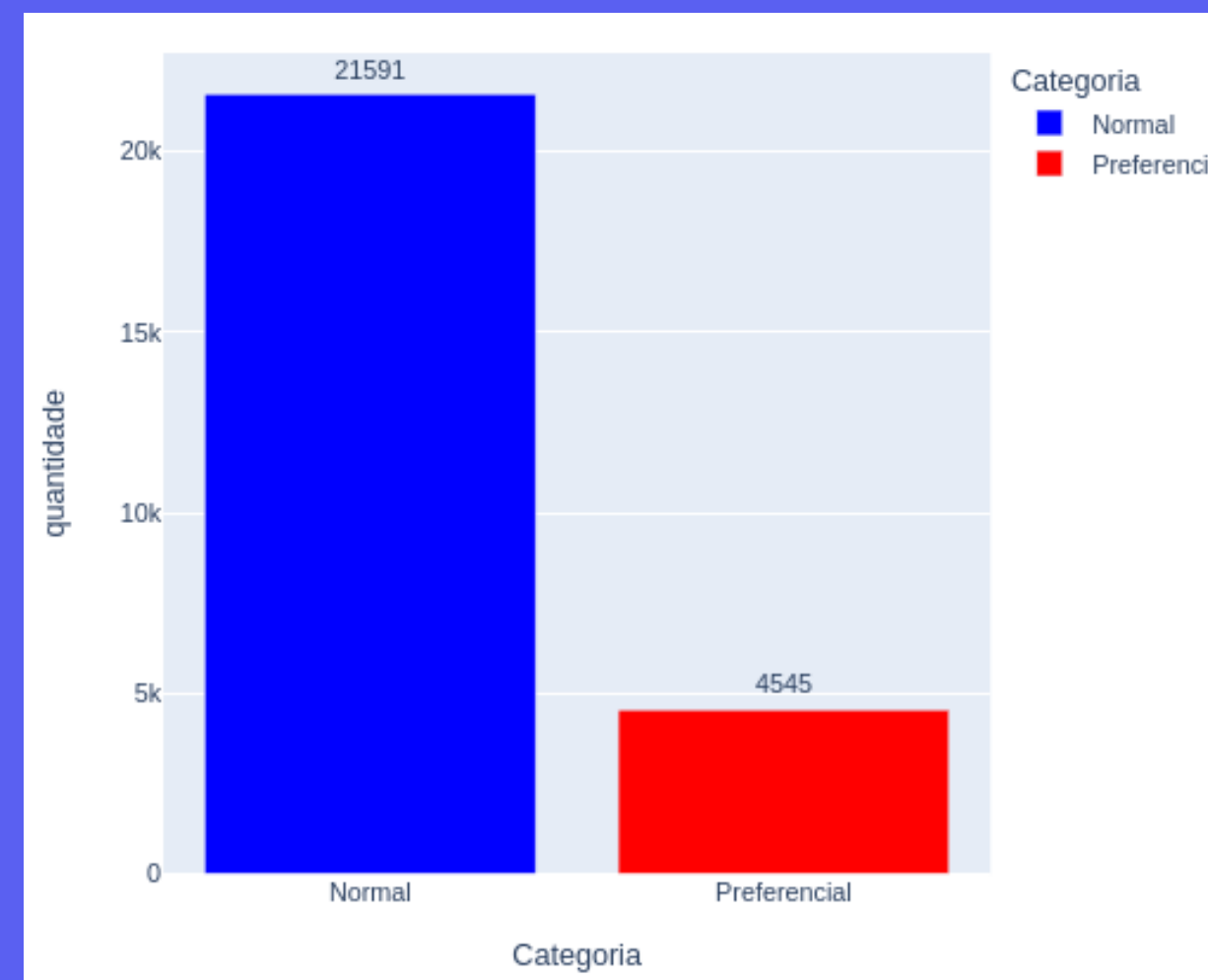
Metodologia



Treinamento dos modelos

Triagem de processos Normais e Preferenciais

➡ "Flag Preferência Legal"



Metodologia



Treinamento dos modelos

Triagem de processos Normais e Preferenciais

- Técnica "Undersampling"
- 3636 documentos para treino
- 4319 documentos da classe "NORMAL"
- 909 da classe "PREFERENCIAL"

Metodologia



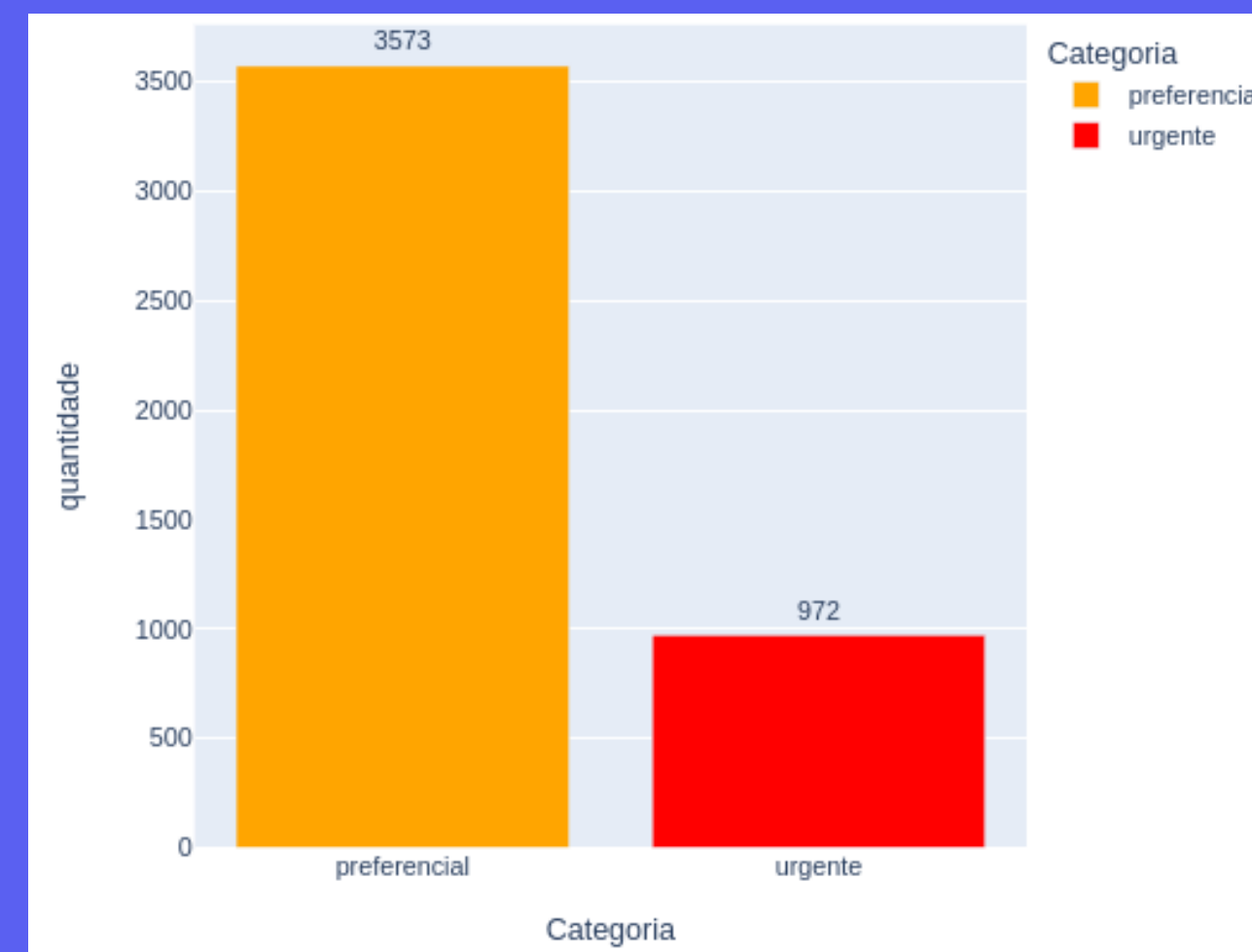
Treinamento dos modelos

Triagem de processos Preferenciais e Urgentes

➡ "Preferência Legal"

URGENTES

- Doença Terminal
- Réu Preso
- Preso/ Acolhido/ Internado



Metodologia

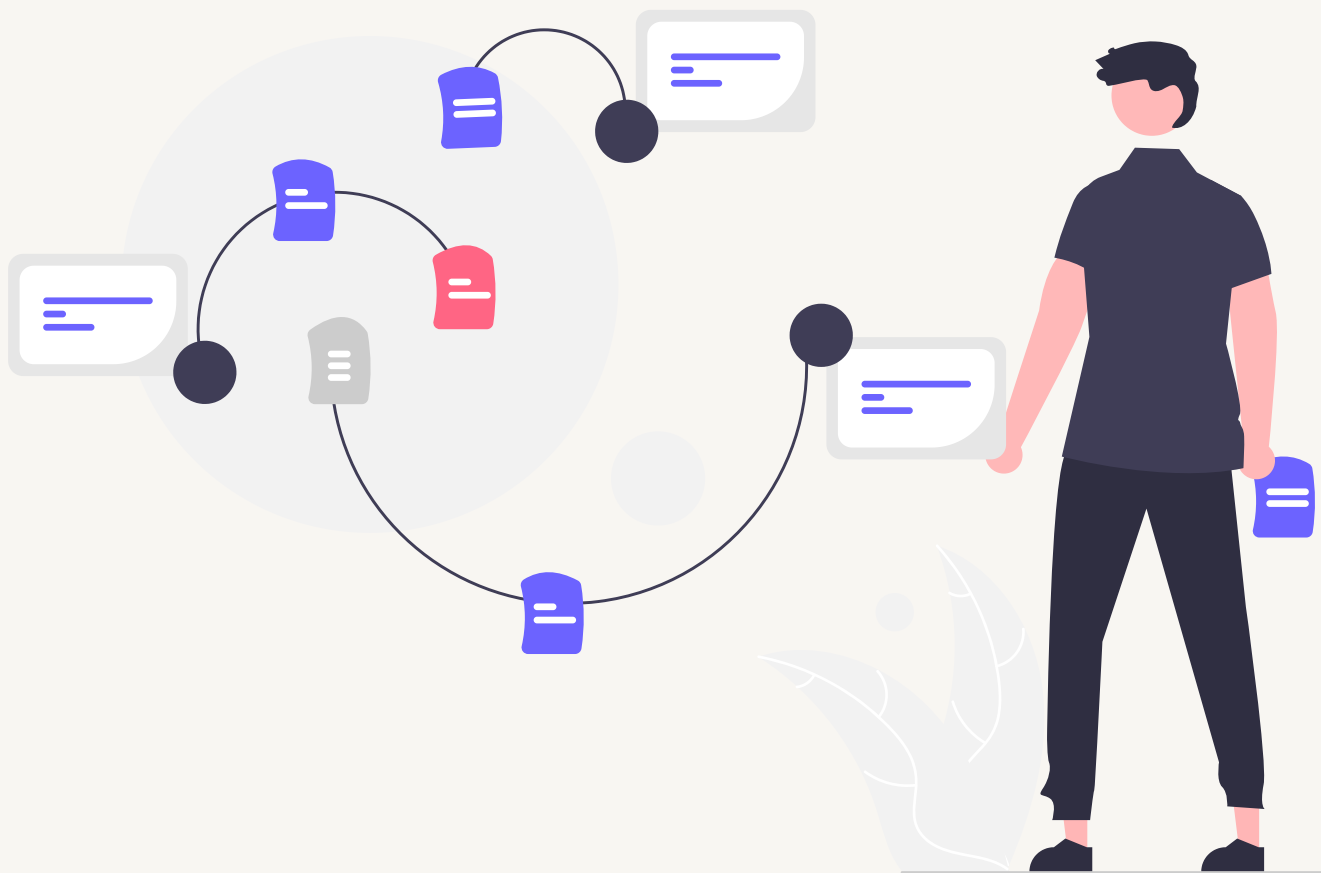


Treinamento dos modelos

Triagem de processos Preferenciais e Urgentes

- Técnica "Undersampling"
- 778 documentos para treino
- 715 documentos da classe "PREFERENCIAL"
- 194 da classe "URGENTE"

Metodologia



Treinamento dos modelos



TF-IDF



ngram_range(1, 3)



max_features=2000



Random Forest

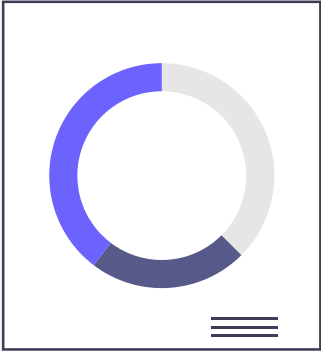
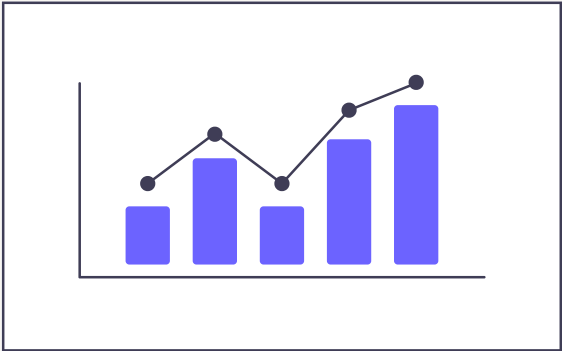


Support Vector Machines (SVM)



ExtraTreesClassifier

Resultados

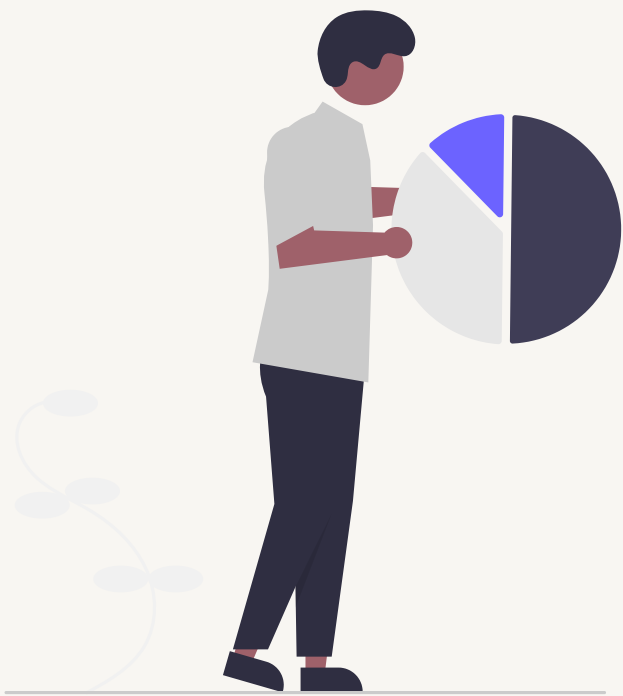


Avaliação dos modelos

Triagem de processos Normais e Preferenciais

| Métricas | Classe | RF | SVM | ExtraTree |
|----------|--------------|-----|-----|-----------|
| Acurácia | | 91% | 92% | 92% |
| Recall | Normal | 92% | 92% | 92% |
| | Preferencial | 91% | 90% | 89% |
| Precisão | Normal | 98% | 98% | 98% |
| | Preferencial | 69% | 70% | 70% |
| F1-Score | Normal | 95% | 95% | 95% |
| | Preferencial | 79% | 79% | 79% |

Resultados



Avaliação dos modelos

Triagem de processos Preferenciais e Urgentes

| Métricas | Classe | RF | SVM | ExtraTree |
|----------|--------------|-----|-----|-----------|
| Acurácia | | 94% | 94% | 94% |
| Recall | Preferencial | 95% | 95% | 95% |
| | Urgente | 90% | 91% | 90% |
| Precisão | Preferencial | 97% | 98% | 97% |
| | Urgente | 83% | 83% | 83% |
| F1-Score | Preferencial | 96% | 96% | 96% |
| | Urgente | 86% | 87% | 87% |

Considerações Finais



Principais contribuições

- Reduzir o tempo de espera da parte autora do processo quando é necessário uma certa urgência
- Economizar tempo dos Servidores da Justiça
- Dar opções de melhorias e contribuir com a transformação digital do Poder Judiciário

Considerações Finais



Limitações

- Quantidade de processos normais que passam pela triagem para a próxima etapa

Considerações Finais



Trabalhos Futuros

- Inclusão de novos dados
- Aprimorar classificação de processos normais
- Testar novos modelos
- Implementar API
- Acoplar a API ao PJe ou a algum projeto de residência

Fim

"A inteligência artificial pode descobrir soluções para problemas do passado, que mudarão a nossa forma de resolver os mesmos problemas no futuro."

(Dawiny Bastos)